1. Sets

$1'1$. Sets and elements

In an intellectual activity, one of the most profound actions is gathering objects in groups. The gathering is performed in mind and is not accompanied with any action in the physical world.

As soon as the group has been created and assigned a name, it can be a subject of thoughts and arguments and, in particular, can be included into other groups.

Mathematics has an elaborated system of notions, which organizes and regulates creating those groups and manipulating them. The system is called the *naive set theory.* This name is a slightly misleading, because this is rather a language than a theory.

The first words in this language are *set* and *element*. By a set we understand an arbitrary collection of various objects. An object included into the collection is an *element* of the set. A set *consists* of its elements. It is also *formed* by them. In order to diversify the wording, the word *set* is replaced by the word *collection*. Sometimes other words, such as *class*, *family*, and *group*, are used in the same sense, but this is not quite safe because each of these words is associated in modern mathematics with a more special meaning, and hence should be used instead of the word *set* with caution.

If x is an element of a set A, then we write $x \in A$ and say that x *belongs to* A and A *contains* x. The sign $\in$ is a variant of the Greek letter epsilon, which corresponds to the first letter of the Latin word *element*. To make the notation more flexible, the formula $x \in A$ is also allowed to be written in the form $A \ni x$. The disrespect to the origin of the notation is payed off by emphasizing a meaningful similarity to the inequality symbols $<$ and $>$. To state that x is not an element of A, we write $x \notin A$ or $A \not\ni x$.

$1'2$. Equality of sets

A set is determined by its elements. The set is nothing but a collection of its elements. This manifests most sharply in the following principle (called *Axiom of Extensionality*):

Two sets are considered equal if and only if they have the same elements.

In this sense, the word *set* has slightly disparaging meaning. When something is called a set, this shows, maybe unintentionally, a lack of interest to whatever organization of the elements of this set.

For example, when we say that a line is a set of points, we assume that two lines coincide if and only if they consist of the same points. On the other hand, we commit ourselves to consider all relations between points on a line (e.g., the distance between points, the order of points on the line, etc.) separately from the notion of line.

We may think of sets as of boxes that can be built effortlessly around elements, just to distinguish them from the rest of the world. The cost of this lightness is that such a box is not more than the collection of elements placed inside. It is a little more than just a name: it is a declaration of our wish to think about this collection of things as of entity and not to go into details about the nature of its members-elements. Elements, in turn, may also be sets, but as long as we consider them elements, they play the role of atoms, with their own original nature ignored.

In modern Mathematics, the words *set* and *element* are very common and appear in most texts. They are even overused, that is used at instances when it is not appropriate. For example, it is not good to use the word *element* as a replacement for other, more meaningful word. When you call something an *element*, then the *set* whose element this one is should be clear. The word *element* makes sense only in combination with the word *set*, unless we deal with a non-mathematical term (like *chemical element*), or a rare old-fashioned exception from the common mathematical terminology (sometimes the expression under the sign of integral is called an *infinitesimal element*; lines, planes, and other geometric images are also called *elements* in old texts). Euclid's famous book on Geometry is called *Elements*, too.

1′3. The empty set

Thus, an element may not be without a set. However, a set may have no elements. Actually, there is such a set. This set is unique because a set is completely determined by its elements. It is the ***empty set*** denoted[1] by $\varnothing$.

1′4. Basic sets of numbers

In addition to $\varnothing$, there are some other sets so important that they have their own special names and denoted by special symbols.

The set of all positive integers, i.e., 1, 2, 3, 4,..., etc., is denoted by $\mathbb{N}$.

The set of all integers, both positive, negative, and the zero, is denoted by $\mathbb{Z}$.

The set of all rational numbers (join to the integers all the numbers that are presented by fractions, like $2/3$ and $\frac{-7}{5}$) is denoted by $\mathbb{Q}$.

[1] Other symbols, like Λ, are also in use, but $\varnothing$ has become most common one.

The set of all real numbers (obtained by adjoining to rational numbers the numbers like $\sqrt{2}$ and $\pi = 3.14\ldots$) is denoted by $\mathbb{R}$.

The set of complex numbers is denoted by $\mathbb{C}$.

1′5. Describing a set by listing its elements

A set presented by a list a, b, ..., x of its elements is denoted by the symbol $\{a, b, \ldots, x\}$. In other words, the list of objects enclosed in curly brackets denotes the set whose elements are listed. For example, $\{1, 2, 123\}$ denotes the set consisting of the numbers 1, 2, and 123. The symbol $\{a, x, A\}$ denotes the set consisting of three elements: a, x, and A, whatever objects these three letters denote.

1.1 What is $\{\varnothing\}$? How many elements does it contain?

1.2 Which of the following formulas are correct:

1) $\varnothing \in \{\varnothing, \{\varnothing\}\}$; 2) $\{\varnothing\} \in \{\{\varnothing\}\}$; 3) $\varnothing \in \{\{\varnothing\}\}$?

A set consisting of a single element is called a *singleton*. This is any set which can be presented as $\{a\}$ for some a.

1.3 Is $\{\{\varnothing\}\}$ a singleton?

Notice that the sets $\{1, 2, 3\}$ and $\{3, 2, 1, 2\}$ are equal since they have the same elements. At first glance, lists with repetitions of elements are never needed. There even arises a temptation to prohibit usage of lists with repetitions in such notation. However, as it often happens to temptations to prohibit something, this would not be wise. Indeed, quite often one cannot say a priori whether there are repetitions or not. For example, the elements in the list may depend on a parameter, and under certain values of the parameter some entries of the list coincide, while for other values they don't.

1.4 How many elements do the following sets contain?

1) $\{1, 2, 1\}$;	2) $\{1, 2, \{1, 2\}\}$;	3) $\{\{2\}\}$;
4) $\{\{1\}, 1\}$;	5) $\{1, \varnothing\}$;	6) $\{\{\varnothing\}, \varnothing\}$;
7) $\{\{\varnothing\}, \{\varnothing\}\}$;	8) $\{x, 3x - 1\}$ for $x \in \mathbb{R}$.	

1′6. Subsets and inclusions

If A and B are sets and every element of A also belongs to B, then we say that A is a ***subset*** of B, or B ***includes*** or ***contains*** A, and write $A \subset B$ or $B \supset A$.

The inclusion signs $\subset$ and $\supset$ resemble the inequality signs $<$ and $>$ for a good reason: in the world of sets, the inclusion signs are obvious counterparts for the signs of inequality. However,there is a deep difference between the notions of inequality and inclusion: no number a satisfies the inequality $a < a$, while any set A contains itself:

1.A Reflexivity of inclusion. Inclusion $A \subset A$ holds true for any A.

Proof. Recall that, by the definition of an inclusion, $A \subset B$ means that each element of A is an element of B. Therefore, the statement that we must prove can be rephrased as follows: each element of A is an element of A. This is tautologically correct. $\square$

Thus, the inclusion signs are not truly genuine counterparts of the inequality signs $<$ and $>$. They are closer to $\leq$ and $\geq$.

Sometimes, being inspired by signs $\leq$ and $\geq$, inclusions are denoted by symbols $\subseteq$ and $\supseteq$ or even $\subseteqq$ and $\supseteqq$, reserving the symbols $\subset$ and $\supset$ for strict inclusions, that prohibit equality, like strict inequalities. We follow the mainstream mathematical notation in which the signs $\subseteq$, $\supseteq$, $\subseteqq$ and $\supseteqq$ are not used and strict inclusions are denoted by $\subsetneq$ and $\supsetneq$ or by $\subsetneqq$ and $\supsetneqq$.

Inclusion $\subset$ and inequality $\leq$ are not only similar, but are closely related. We will discuss this later, section **??**.

1.B Ubiguity of the empty set. $\varnothing \subset A$ for any set A. In other words, the empty set is present in each set as a subset.

Proof. Recall that, by the definition of inclusion, $A \subset B$ means that each element of A is an element of B. Thus, we need to prove that any element of $\varnothing$ belongs to A. This is true because $\varnothing$ does not contain any elements. $\square$

It may happen that you are not satisfied with this proof. Arguments about the empty set may confuse at first. To this end, look at

Another proof of 1.B. Let us resort to the question whether the statement which we prove can be wrong. How can it happen that $\varnothing$ is not a subset of A? This is possible only if $\varnothing$ contains an element which is not an element of A. However, $\varnothing$ does not contain such elements because $\varnothing$ contains no elements at all. $\square$

Thus, each set A has two obvious subsets: the empty set $\varnothing$ and A itself. A subset of A different from $\varnothing$ and A is a ***proper*** subset of A. This word is used when we do not want to consider the obvious subsets (which are ***improper***).

1.C Transitivity of inclusion. If A, B, and C are sets, $A \subset B$ and $B \subset C$, then $A \subset C$.

Proof. We must prove that each element of A is an element of C. Let $x \in A$. Since $A \subset B$, it follows that $x \in B$. Since $B \subset C$, the latter (i.e., $x \in B$) implies $x \in C$. This is what we had to prove. □

1′7. Defining a set by a condition (a set-builder notation)

As we know (see 1′5), a set can be described by presenting a list of its elements. This simplest way may be not available or, at least, be not the easiest one. For example, it is easy to say: "the set of all solutions of the following equation" and write down the equation. This is a reasonable description of the set. At least, it is unambiguous. Having accepted it, we may start speaking on the set, studying its properties, and eventually may be lucky to solve the equation and obtain the list of its solutions. Although the latter task may be difficult, this should not prevent us from discussing the set until the time when the equation will be solved. (Solution of some equations took centuries!)

Thus, we see another way for describing a set: formulate properties that distinguish the elements of the set among elements of some wider and already known set. Here is the corresponding notation:

The subset of a set A consisting of the elements x that satisfy a condition $P(x)$ is denoted by $\{x \in A \mid P(x)\}$.

1.5 Present the following sets by lists of their elements (i.e., in the form $\{a, b, \dots\}$)
(a) $\{x \in \mathbb{N} \mid x < 5\}$, (b) $\{x \in \mathbb{N} \mid x < 0\}$, (c) $\{x \in \mathbb{Z} \mid x < 0\}$.

The set-builder notation unveils a close relation between logic statements and sets. Every statement P about elements of a set A defines a subset $\{x \in A \mid P(x)\}$ of A. On the other hand, any subset $B \subset A$ gives rise to a property of elements of A: namely, the property of belonging to B, that is $x \in B$.

For example, let us figure out what on the side of logic statements corresponds to inclusion. Let B and C are subsets of a set A. Let $B = \{x \in A \mid P(x)\}$ and $C = \{x \in A \mid Q(x)\}$, that is P and Q are the statements defining B and C, respectively. Inclusion $B \subset C$ means that each element of B is an element of C. In other words, if $x \in B$, then $x \in C$, or, in terms of P and Q, if $P(x)$, then $Q(x)$. Thus, the inclusion $B \subset C$ corresponds to implication "if $P(x)$, then $Q(x)$".

1′8. Conditional and biconditional

In everyday English the meaning of the "if ... then" construction is ambiguous. The construction "if P, then Q" always means that if P is true, then Q is true also. Sometimes that is all it means; other times it means something more: that if P is false, Q must be false either. In the first case we say about *conditional statement*, in the second case, about *biconditional*. In ordinary everyday English, usually one decides from the context whether conditional or biconditional sense is intended.

Mathematicians tend to avoid ambiguities. They have agreed to use the construction "if ... then" in the first, conditional sense, as above, so that a statement of the form "If P, then Q" means that if P is true, Q is true also, but if P is false, Q may be either true or false.

There is an important exception from this agreement: in a definition, when a new word is introduced, mathematicians use "if" in the biconditional sense. For example, when defining the notion of subset, we say: "A is a subset of B if each element of A belongs to B" - and this means that whenever we say that A is a subset of B, each element of A does belong to B. An extra evidence that the word "if" is understood biconditionally are expressions *"is called"*, *"is said to be"* and *"one says that"*, which introduce new words.

However, this is the only exception. In any other mathematical context a sentence "P if Q" is *not* understood biconditionally.

Biconditional statements are not rare guests in the mathematical language. They appear very often, probably, more often than in everyday English. They are presented by the words *"if and only if*. In writing, this expression is often abbreviated to a single word *iff*. So, we write "P iff Q" instead of "P if and only if Q. Another way to express the same: "P is *necessary and sufficient* for Q." There is also a formula-synonym: $P \iff Q$.

A conditional sentence "if P, then Q" also can be replaced by formula: $P \Rightarrow Q$. Here is a list of other ways to say the same:

- P is sufficient for Q,
- Q is necessary for P,
- P only if Q,
- P implies Q.

1′9. For proving equality of sets, prove two inclusions

Working with sets, we need from time to time to prove that two sets, say A and B, which may have emerged in quite different ways, are equal. The most common way to do this is provided by the following theorem.

1.D Test for equality of sets.

$$A = B \text{ if and only if } A \subset B \text{ and } B \subset A.$$

Proof. We have already seen that $A \subset A$. Hence, if $A = B$, then, indeed, $A \subset B$ and $B \subset A$. On the other hand, $A \subset B$ means that each element of A belongs to B, while $B \subset A$ means that each element of B belongs to A. Hence, A and B have the same elements, i.e., they are equal. □

1′10. Inclusion versus belonging

1.E $x \in A$ *if and only if* $\{x\} \subset A$.

Despite this obvious relation between the notions of belonging $\in$ and inclusion $\subset$ and similarity of the symbols $\in$ and $\subset$, the concepts are quite different. Indeed, $A \in B$ means that A is an element in B (i.e., one of the indivisible pieces constituting B), while $A \subset B$ means that A is made of some of the elements of B.

In particular, we have $A \subset A$, while $A \notin A$ for any reasonable A. Thus, *belonging is not reflexive*. One more difference: *belonging is not transitive*, while inclusion is.

1.F Non-transitivity of belonging. Construct three sets A, B, and C such that $A \in B$ and $B \in C$, but $A \notin C$. Cf. 1.C.

Construction. Take $A = \{1\}$, $B = \{\{1\}\}$, and $C = \{\{\{1\}\}\}$. □

Remark. It is more difficult to construct sets A, B, and C such that $A \in B$, $B \in C$, and $A \in C$. Though, it is possible. Take, for example, $A = \{1\}$, $B = \{\{1\}\}$, and $C = \{\{1\}, \{\{1\}\}\}$.

1.G Non-reflexivity of belonging. Construct a set A such that $A \notin A$. Cf. 1.A.

Construction. It is easy to construct a set A with $A \notin A$. Take $A = \varnothing$, or $A = \mathbb{N}$, or $A = \{1\}, \dots$ □

1.6 May belonging be reflexive for a set? Construct a set X such that $X \in X$.

Construction. A set X such that $X \in X$ is a strange creature. It would not appear in a real problem, unless you think really globally. Take for X the set of all sets. □

Mathematicians avoid such sets. There are good reasons for this. If we think overly globally, the thoughts may become insane. If we consider the set of all sets, then why not to consider the set Y of all sets X such that $X \notin X$? Does Y belongs to itself? If $Y \in Y$, then $Y \notin Y$ since each element X of Y has the property that $X \notin X$. If $Y \notin Y$, then $Y \in Y$ since Y is the set of ALL sets X such that $X \notin X$. This contradiction shows that our definition of Y does not make sense. An easy way to avoid this paradox is to prohibit consideration of sets with the property $X \in X$. The set of all sets is not a legitimate set.

1′11. Intersection and union

The *intersection* of sets A and B is the set formed of their common elements, i.e., elements belonging both to A and B. It is denoted by $A \cap B$ and described by the formula

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$

Sets A and B are said to be *disjoint* if $A \cap B = \varnothing$. In other words, sets are disjoint if they have no common elements.

The *union* of sets A and B is
the set formed by all elements that belong to at least one of the two sets.
The union of A and B is denoted by $A \cup B$. It is described by the formula

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}.$$

Here the conjunction *or* should be understood in the inclusive way: the statement "$x \in A$ or $x \in B$" means that x belongs to *at least one* of the sets A and B, and, maybe, to both of them. This agrees with the usage of the word "or" commonly accepted in mathematics.

In everyday English, the word "or" is ambiguous. Sometimes the statement "P or Q" means "P or Q, or both" and sometimes it means "P or Q, but not both". The intended meaning usually is recovered from the context.

Mathematicians tend to keep their language free of ambiguities. In particular, they have agreed to use the word "or" only in the first sense, so that the statement "P or Q" always means "P or Q, or both." If one means "P or Q, but not both," then one has to include the phrase "but not both" explicitly.

1.H Commutativity of $\cap$ and $\cup$. For any two sets A and B, we have

$$A \cap B = B \cap A \quad \text{and} \quad A \cup B = B \cup A.$$

1.7 Prove that for any set A we have

$$A \cap A = A, \qquad A \cup A = A, \qquad A \cup \varnothing = A, \quad \text{and} \quad A \cap \varnothing = \varnothing.$$

1.8 Prove that for any sets A and B we have[2]

$$A \subset B, \quad \text{iff} \quad A \cap B = A, \quad \text{iff} \quad A \cup B = B.$$

1.I Associativity of $\cap$ and $\cup$. For any sets A, B, and C, we have

$$(A \cap B) \cap C = A \cap (B \cap C) \quad \text{and} \quad (A \cup B) \cup C = A \cup (B \cup C).$$

Associativity allows us not to care about brackets and sometimes even omit them. We define $A \cap B \cap C = (A \cap B) \cap C = A \cap (B \cap C)$ and $A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C)$.

1′12. The notion of map

A *map* f of a set X to a set Y is a triple consisting of X, Y, and a rule,[3] which assigns to every element of X exactly one element of Y.

There are other words with the same meaning: *mapping*, *function*, etc. (Special kinds of maps may have special names like *functional*, *operator*, *sequence*, *family*, *fibration*, etc.)

If f is a map of X to Y, then we write $f : X \to Y$, or $X \xrightarrow{f} Y$. The element b of Y assigned by f to an element a of X is denoted by $f(a)$ and called the *image* of a under f, or the *f-image* of a. In order to state that $b = f(a)$, one may write also $a \overset{f}{\mapsto} b$, or $f : a \mapsto b$. We also define maps by formulas like $f : X \to Y : a \mapsto b$, where b is explicitly expressed in terms of a.

1.9 Let X and Y be sets consisting of p and q elements, respectively. Find the number of maps $X \to Y$.

1′13. The main classes of maps

A map $f : X \to Y$ is a *surjective map*, or just a *surjection* if every element of Y is the image of at least one element of X. (We also say that f is *onto*.) A map $f : X \to Y$ is an *injective map*, *injection*, or *one-to-one map* if every element of Y is the image of at most one element of X. A map is a *bijective map*, *bijection* if it is both surjective and injective.

1.10 Let X and Y be sets consisting of p and q elements, respectively. Find the number of injections $X \to Y$.

1.11* *Let X and Y be sets consisting of p and q elements, respectively. Find the number of surjections $X \to Y$.*

[2] Here, as usual, *iff* stands for "if and only if".

[3] Certainly, the rule (as everything in the set theory) may be thought of as a set. Section ?? below. Namely, the rule can be converted to (or, if you prefer, encoded by) the set Γ_f of the ordered pairs (x, y) with $x \in X$ and $y \in Y$ such that the rule assigns y to x. This is the *graph* of f. It is a subset of $X \times Y$. Recall that $X \times Y$ is the set of all ordered pairs (x, y) with $x \in X$ and $y \in Y$.

1′14. Identity and inclusion

The *identity map* of a set X is the map $\mathrm{id}_X : X \to X : x \mapsto x$. It is denoted by id if X is clear from the context. If A is a subset of X, then the map $\mathrm{in}_A : A \to X : x \mapsto x$ is the *inclusion map*, or just *inclusion*, of A into X. It is denoted just by in when A and X are clear.

1′15. Composition

The *composition* of maps $f : X \to Y$ and $g : Y \to Z$ is the map $g \circ f : X \to Z : x \mapsto g\big(f(x)\big)$.

1.J Associativity. *$h\circ(g\circ f) = (h\circ g)\circ f$ for any maps $f : X \to Y$, $g : Y \to Z$, and $h : Z \to U$.*

Proof. Let $x \in X$. Then

$$h \circ (g \circ f)(x) = h(g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = (h \circ g) \circ f(x).$$

1.K *$f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f$ for any $f : X \to Y$.*

1.L *A composition of injections is injective.*

Proof. Let $x_1 \neq x_2$. Then $f(x_1) \neq f(x_2)$ because f is injective, and $g(f(x_1)) \neq g(f(x_2))$ because g is injective.

1.M *If the composition $g \circ f$ is injective, then so is f.*

Proof. If f is not injective, then there exist $x_1 \neq x_2$ with $f(x_1) = f(x_2)$. However, then $(g \circ f)(x_1) = (g \circ f)(x_2)$, which contradicts the injectivity of $g \circ f$.

1.N *A composition of surjections is surjective.*

Proof. Let $f : X \to Y$ and $g : Y \to Z$ be surjective. Then we have $f(X) = Y$, whence $g(f(X)) = g(Y) = Z$.

1.O *If the composition $g \circ f$ is surjective, then so is g.*

Proof. This follows from the obvious inclusion $\mathrm{Im}(g \circ f) \subset \mathrm{Im}\, g$.

1.P *A composition of bijections is a bijection.*

Proof. This follows from 1.L and 1.N.

1.12 Let a composition $g \circ f$ be bijective. Is then f or g necessarily bijective?

1′16. Inverse and invertible

A map $g : Y \to X$ is *inverse* to a map $f : X \to Y$ if $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. A map having an inverse map is *invertible*.

1.Q *A map is invertible iff it is a bijection.*

Proof. $\boxed{\Longrightarrow}$ Use 1.M and 1.O. $\boxed{\Longleftarrow}$ Let $f : X \to Y$ be a bijection. Then, by the surjectivity, for each $y \in Y$ there exists $x \in X$ such that $y = f(x)$, and, by the injectivity, such an element of X is unique. Putting $g(y) = x$, we obtain a map $g : Y \to X$. It is easy to check (please, do it!) that g is inverse to f.

1.R *If an inverse map exists, then it is unique.*

Proof. This is actually obvious. On the other hand, it is interesting to look at a "mechanical" proof. Let two maps $g, h : Y \to X$ be inverse to a map $f : X \to Y$. Consider the composition $g \circ f \circ h : Y \to X$. On the one hand, we have $g \circ f \circ h = (g \circ f) \circ h = \mathrm{id}_X \circ h = h$. On the other hand, we have $g \circ f \circ h = g \circ (f \circ h) = g \circ \mathrm{id}_Y = g$.

2. Numbers

2′1. Natural numbers

For most people, mathematics starts with numbers. Everybody knows natural numbers. They are 1, 2, 3, 4, 5, ... Natural numbers are used to count the number of elements in finite sets. The nature of the set and its elements does not matter.

Natural numbers come with several structures. They are ordered: $1 < 2 < 3 < 4 < \dots$. Arithmetic operations can be performed on natural numbers. Natural numbers come with a whole package of these things. That's what we mean speaking about the *system* of natural numbers.

Then other systems of numbers come, because natural numbers are insufficient for many purposes. Some objects that we count can be split into pieces, and to take the pieces into account, one need to extend the notion of number and introduce fractions. Then similarly one needs to add negative numbers. Then one realizes that some quantities are not co-measurable and various sorts of real numbers appear. Then one realizes that real numbers do not allow us to solve some equations, and we add complex numbers. Then... This process continues further, although the school frameworks do not reach that far.

The beginning of the process hides in the memory of childhood. In these notes, it is revisited. We start with the natural numbers. The goal is to show the roots and reasonings that lie behind the notions which are commonly known.

Natural numbers are so close to the basics of the human nature that most mathematicians prefer to consider them given by God. Indeed, there is a famous phrase attributed to a German mathematician Leopold Kronecker: "God created the natural numbers, and all the rest is the work of man." The other numbers, like integers, rational numbers, real numbers, complex numbers, etc. can be built up of the natural numbers given by God.

Leopold Kronecker (1823-1891)

The natural numbers are important habitants of the mathematical universe. The set of all natural numbers is denoted by a special symbol $\mathbb{N}$. This is a rare honor for an object in the mathematical world to possess a special commonly accepted notation.

As usual, what is claimed to be given by God can be explained without him.

2′2. Cardinal numbers

Numbers can be built up out of the task that they serve for. A natural number is used to measure the number of elements in a set. Thus a natural number can be described via the sets that have this number of elements.

Sets with the same number of elements admit a bijection of one to the other.

2.A *Existence of bijections between sets is an equivalence relation.*

Observe that this theorem is true for any sets (not only for finite ones). Let us formulate it more explicitly, with extra details. These extra details really prove it.

- *For each set X there exists a bijection $X \to X$.* In particular, the identity map $\mathrm{id} : X \to X$ (which maps each $a \in X$ to the same a) is a bijection.
- *If there exists a bijection $f : X \to Y$, then there exists a bijection $Y \to X$.* In particular, the inverse map $f^{-1} : Y \to X$ is a bijection.
- *If there exist bijections $f : X \to Y$ and $g : Y \to Z$, then there exists a bijection $X \to Z$.* In particular, the composition $g \circ f : X \to Z$ is a bijection.

These statements mean, respectively, reflexivity, symmetry, and transitivity of the relation "*there exists a bijection $X \to Y$.*" Thus this relation is an *equivalence relation*. Sets X and Y such that there exists a bijection $X \to Y$ are said to be *equinumerous* or *equipotent*.

An equivalence relation defines a partition of all sets into classes of equivalent sets (called also equivalence classes). Equivalence classes for the relation of being equinumerous are called *cardinal numbers*. The cardinal number of a set X is denoted by $\operatorname{card}(X)$. This is the class of all sets that admit bijections to X (and hence to each other).

A set can be either finite or infinite. Natural numbers are the cardinal numbers of finite sets. Although our main interest is the system of natural numbers, for a while we will study a broader system of all cardinal numbers. We will restrict ourselves to natural numbers as soon as we will come across the properties that are not shared with all cardinal numbers.

2′3. Operations with sets and numbers

Arithmetic operations with numbers come from the corresponding operations with finite sets.

Addition is related to union: the number of elements in the union of sets A and B is the sum of the numbers of elements in A and in B, provided A and B are disjoint (i.e., $A \cap B = \varnothing$).

Similarly, multiplication of numbers is related to the corresponding operation with sets. The (*Cartesian*) *product* of sets A and B is the set of ordered pairs (a, b) with the first element a taken from A and the second element b taken from B:

$$A \times B = \{(a, b) \mid a \in A,\ b \in B\}.$$

The number of $A \times B$ is the product of the numbers of elements in A and in B.

The union and Cartesian product operations are defined for arbitrary sets, no matter finite or infinite. This motivates the following generalizations of addition and multiplication from natural numbers to arbitrary cardinal numbers. Most important properties are preserved under the generalizations. This gives an opportunity to revisit the properties and discuss their nature and proofs in the new high generality.

Let a and b be cardinal numbers. Their sum $a + b$ is defined as the cardinal number of the union $A \cup B$ of any two disjoint sets A and B that have cardinal numbers a and b, respectively.

This definition requires a proof. Indeed, it presumes that the cardinal number of $A \cup B$ does not depend on choice of A and B. Let us prove that this is really the case.

Let A' and B' be other disjoint sets with cardinal numbers a and b. Then there exist bijections $f : A \to A'$ and $g : B \to B'$, and they define a map $A \cup B \to A' \cup B'$ which coincides with f on A and with g on B. Since $A \cap B = \varnothing$, this map is well-defined. (If the intersection was not empty, there could be ambiguity for a common element of A and B: maps f and g could map it to different elements.) The map $A \cup B \to A' \cup B'$ defined by f and g is bijective, since f and g are bijective. □

The operation of union is associative and commutative, therefore the addition of cardinal numbers is associative and commutative. (Recall that associativity means $(a + b) + c = a + (b + c)$, and commutativity means $a + b = b + a$.)

Multiplication of cardinal numbers is defined similarly: the product $a \cdot b$ of cardinal numbers a and b is defined as the cardinal number of the Cartesian product $A \times B$ of sets A and B that have cardinal numbers a and b, respectively.

Like the definition of sum of cardinal numbers above, this definition requires a proof of independence on the choice of A and B. Let A' and B' be other sets with cardinal numbers a and b and let $f : A \to A'$ and $g : B \to B'$ be bijections. Then one can build up a bijection $A \times B \to A' \times B'$. Namely let us map (a, b) to $(f(a), g(b))$. This map is invertible, the inverse to it is made out of the maps inverse to f and g in the same manner.

Multiplication of cardinal numbers has the same properties as the multiplication of natural numbers.

2.B . Commutativity of multiplication.. *For any cardinal numbers a and b,*

$$a \cdot b = b \cdot a$$

Proof. Let A and B be sets with cardinal numbers a and b, respectively. Then $A \times B$ has cardinal number $a \cdot b$ and $B \times A$ has cardinal number $b \cdot a$. The equality between these cardinal numbers follows from existence of the bijection $A \times B \to B \times A$. Such a bijection is defined by the formula $(a, b) \mapsto (b, a)$. □

2.C . Associativity of multiplication.. *For any cardinal numbers a, b and c,*

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

Proof. Let A, B and C be sets with cardinal numbers a, b and c, respectively. The theorem is proved by constructing a bijection $(A \times B) \times C \to A \times (B \times C)$. This is done by formula $((a, b), c) \mapsto (a, (b, c))$. This is nothing but a bare rearrangement of parantheses. □

2.D . Distributivity.. *For any cardinal numbers a, b and c,*

$$(a + b) \cdot c = a \cdot c + b \cdot c.$$

Proof. Let A, B and C be sets with cardinal numbers a, b and c, respectively. Assume that $A \cap B = \varnothing$. Then the set $(A \cup B) \times C$ has cardinal number $(a + b) \cdot c$. The sets $A \times C$

and $B \times C$ are disjoint, the set $(A \times C) \cup (B \times C)$ has cardinal number $a \cdot c + b \cdot c$ and $(A \cup B) \times C = (A \times C) \cup (B \times C)$. □

2′4. The zero

The cardinal number of $\varnothing$ is called the *zero* and denoted by 0. Sometimes it is considered a natural number, sometimes not. Historically it appeared much later than other natural numbers.

The empty set is disjoint with any set A. Obviously, $A \cup \varnothing = A$ for any set A. Thus $a + 0 = a$ for any cardinal number a.

What is $a \cdot 0$? In order to answer to this question, consider an arbitrary set A and its product by the empty set. The set $A \times \varnothing$ consists of pairs (x, y) where x in an element of A and y is an element of $\varnothing$. However, $\varnothing$ has no element. Thus $A \times \varnothing = \varnothing$ and $a \cdot 0 = 0$ for any cardinal number a.

2′5. The one

A set consisting of a single element is called a *singleton*. For whatever object A one can construct the set that consists just of this object. It is denoted by $\{A\}$. The object A may be a set itself, and may contain lots of elements, or may have no elements at all (this happens if $A = \varnothing$), but speaking about $\{A\}$ we keep in mind only one thing, as an element: A. The set $\{\varnothing\}$ is not empty, it is a singleton, it has the only element, and this element is the empty set $\varnothing$.

Between two singletons X and Y only one map exists: the only element of X is mapped to the only element of Y. In other words, the set of all maps from a singleton to a singleton is a singleton.

Notice that the only map from a singleton to a singleton is a bijection. Thus all the singletons are equinumerous. The cardinal number represented by them is called one and denoted by 1.

2.E *For any cardinal number a, the product $a \cdot 1$ equals a.*

Proof. Let A be a set with $\mathrm{card}(A) = a$ and B be a singleton, $B = \{y\}$. Then $\mathrm{card}(A \times B) = a \cdot 1$. Bijection $A \times B \to A : (x, y) \mapsto x$ proves the equality $a \cdot 1 = a$. □

2′6. Ordering of cardinal numbers

2.F *Let X, Y be sets and $X \neq \varnothing$. An injection $X \to Y$ exists if and only if a surjection $Y \to X$ exists.*

Proof. If $\varphi : X \to Y$ is an injection. Consider a map $\psi : X \to \varphi(X)$ defined by φ. It maps each $a \in X$ to $\varphi(a)$. The map ψ is both injection (as a submap of injection) and surjection. Hence, it is a bijection, and is invertible. Take the inverse map $\psi^{-1} : \varphi(X) \to X$ and extend it somehow to the whole $Y \supset \varphi(X)$. Since $X \neq \varnothing$, there exists an element $a \in X$, and we can map to a each element b of Y which does not belong to $\varphi(X)$. The extended map $Y \to X$ is surjective, because even its restriction to $\varphi(X)$ is surjective.

Now assume that there is a surjective map $\psi : Y \to X$. For each $a \in X$, its preimage $\psi^{-1}(a) = \{b \in Y \mid \psi(b) = a\}$ is not empty (becuase of surjectivity of ψ). Choose for each $a \in X$ an element from the $\psi^{-1}(a)$ and denote it by $\phi(a)$. This defines a map $\phi : X \to Y$. It is injective, because for any two distinct $a, a' \in X$ the preimages $\psi^{-1}(a)$ and $\psi^{-1}(a')$, from which we selected $\phi(a)$ and $\phi(a')$, are disjoint. □

Observe that we used the assumption $X \neq \varnothing$ only in the first part of the proof, when we proved that existence of injection implies existence of surjection in the opposite direction. The assumption is indeed necessary, because the empty set is mapped injectively into any set, while a non-epmty set does not admit any map into the empty one.

A cardinal number a is said to be ***not greater*** than a cardinal number b if for a set A with cardinal number a and B with cardinal number b there exists an injection $A \to B$. The statement a is not greater than b is expressed by formula $a \leq b$ (the same formula as for usual numbers).

2.1 Prove that this definition is correct: the existence of injection does not depend on the choice of representatives A and B of a and b.

Notice that this inequality is not strict: an injection $A \to B$ may happen to be a bijection, and existence of bijection ensures that $a = b$.

Moreover, existence on injection which is not a bijection does not contradict to the equality. Indeed, there exists a non-bijective injections $\mathbb{N} \to \mathbb{N}$, for example, $\mathbb{N} \to \mathbb{N} : n \mapsto n + 1$.

The inequality for cardinal numbers has the usual properties of non-strict inequality:

- **Reflexivity.** $a \leq a$ for any cardinal number a;
- **Transitivity.** if $a \leq b$ and $b \leq c$, then $a \leq c$ for any cardinal numbers a, b, c;
- **Totality.** $a \leq b$ or $b \leq a$ for any cardinal numbers a, b;
- **Antisymmetry.** if $a \leq b$ and $b \leq a$, then $a = b$.

The last property, antisymmetry, is well-known as Cantor-Bernstein-Schroeder theorem. There is a nice elementary proof, see wikipedia. The others have easier proofs, that are left to the reader as exercises.

Overall, the part of theory of cardinal numbers outlined above provides a clear and elementary foundation for the standard order in the set of natural numbers. However, there are specific properties of finite cardinal numbers (i.e., natural numbers), that force us to stay with natural numbers.

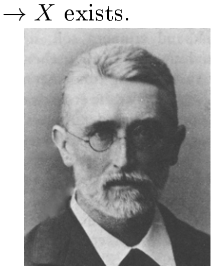
2′7. Special properties of finite sets and numbers

What set is finite? Intuitive idea of finite set is that its elements can be enumerated by finitely many natural numbers.

There are properties distinguishing finite sets from infinite ones and formulated without use of numbers. Probably the most convenient of them is this one:

2.G (Dedekind). *A set X is finite if and only if there exists no injective but not surjective map $X \to X$.*

Proof. Assume that an injective non-surjective map $f : X \to X$ exists. Then $X \smallsetminus f(X) \neq \varnothing$. It is mapped by f bijectively onto $f(X \smallsetminus f(X)) = f(X) \smallsetminus f(f(X))$, which in turn is mapped by f bijectively onto $f(f(X)) \smallsetminus f(f(f(X)))$, etc. Notice that these sets are pairwise disjoint and non-empty. Take an element of the first of them, $a \in X$ such that $a \notin f(X)$. It gives rise to an infinite sequence $a, f(a), f(f(a)), f(f(f(a))), \dots$. Thus, X is infinite as it contains an infinite set.



Richard Dedekind (1831-1916)

Assume that X is infinite. Then it contains an infinite sequence $x_1, x_2, x_3, \dots$ of pairwise distinct elements. Construct the map that maps x_i to x_{i+1} and is identity on the complement of the $\{x_1, x_2, x_3, \dots\}$. This map is injection, but x_1 does not belong to its image, therefore it is not bijective. □

2.H Corollary. *A set X is finite if and only if there exists no surjective non-bijective map $X \to X$.*

Proof. It follows from theorems 2.G and 2.F. □

2.I Corollary. *$a = a + 1$ for any infinite cardinal number a.*

Proof. First, notice that $a \leq a + 1$. On the other hand, let A be a set with $\operatorname{card} A = a$. By Theorem 2.G, there is an injective but not surjective map $f : A \to A$. Let $\{b\}$ be a singleton with $b \notin A$. Let $a \in A \smallsetminus f(A)$. Define a map $A \cup \{b\} \to A$ as f on A and sending b to a. It is injective. So $a + 1 \leq a$. Hence $a + 1 = a$. □

2′8. Making natural numbers out of nothing

The empty set allows to generate some simple non-empty sets. For example, the set $\{\varnothing\}$ consists of a single element, which is the empty set $\varnothing$. The cardinal number of $\{\varnothing\}$ is called *one* and denoted by 1.

Then one can create a set $\{\varnothing, \{\varnothing\}\}$ consisting of two elements: $\varnothing$ and $\{\varnothing\}$. The cardinal number of this set is called *two* and denoted by 2.

And so on... Each next set is obtained by adjoining to the preceding one a single new element. In the desertous environment of initial notions of the set theory for this extra element we may take just the preceding set, which was just cooked up in the preceding step.

The first sets in the sequence looks as follows:

$$\begin{aligned}
&\{\varnothing\},\\
&\{\varnothing, \{\varnothing\}\},\\
&\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\},\\
&\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\},\\
&\dots
\end{aligned}$$

More generally, for every set X we define the *successor* X^+ of X to be the set obtained by adjoining X to the elements of X; in other words $X^+ = X \cup \{X\}$. The same notation will be used for the cardinal numbers: if x denotes the cardinal number of a set X, then x^+ denotes the cardinal number of the set X^+. Then the (canonical sets representing) natural numbers are defined as follows.

$$\begin{aligned}
0 =& \operatorname{card}(\varnothing),\\
1 =& 0^+ = \operatorname{card}(\varnothing^+) = \operatorname{card}(\{\varnothing\}),\\
2 =& 1^+ = \operatorname{card}(\{\varnothing\}^+) \operatorname{card}(\{\varnothing, \{\varnothing\}\}),\\
3 =& 2^+ = \operatorname{card}(\{\varnothing, \{\varnothing\}\}^+) = \operatorname{card}\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\},\\
&\dots
\end{aligned}$$

The cardinal numbers $1, 2, 3, \dots$ obtained in this construction, can be identified with the natural numbers sharing their names.

Notice that the set of numbers $\{0, 1, 2, \dots, n\}$ has cardinal number n^+. So, the construction above can be also described by formula

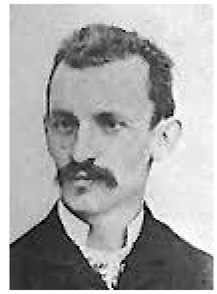
$$n^+ = \operatorname{card}(\{0, 1, 2, \dots, n\}).$$

Observe also that $n^+ = n + 1$ for any cardinal number n.

2′9. Axioms for natural numbers

The natural numbers line up in the order of increase. It starts from number 1. Then for each natural number n there is the next one, the number $n + 1$, the successor of n which is greater than than n by one. It differs from all the preceding numbers.

This well-known image is formalized by Peano axioms. In 1889 a young Italian mathematician Giuseppe Peano published a collection of five axioms in his book, *The principles of arithmetic presented by a new method* (Latin: *Arithmetices principia, nova methodo exposita*).



Giuseppe Peano (1858-1932) in 1887.

Peano axiom 1. There is number $1 \in \mathbb{N}$.

Peano axiom 2. For any $n \in \mathbb{N}$ there exists a unique $n' \in \mathbb{N}$, called the *successor* of n.

Peano axiom 3. For all $n \in \mathbb{N}$, $n' \neq 1$.

Peano axiom 4. Given elements n and m in $\mathbb{N}$, if $n' = m'$ then $n = m$.

Peano axiom 5. If K is a subset of $\mathbb{N}$ such that $1 \in K$, and together with each element n of K the successor n' of n is also contained in K, then $K = \mathbb{N}$.

In the contemporary mathematical language these axioms describe $\mathbb{N}$ as

- a set containing an element 1 (this is Peano axiom 1)
- and equipped with a map $n \mapsto n'$ (this is Peano axiom 2)
- which is injective (Peano axiom 3)
- and such that the image of this map is contained in $\mathbb{N} \setminus \{1\}$ (Peano axiom 4)
- and the only subset of $\mathbb{N}$ containing 1 and invariant under the map $n \mapsto n'$ is the whole $\mathbb{N}$. (Peano axiom 5).

Peano axiom 5 is related to proofs by mathematical induction. Let someone want to prove for each n a statement $S(n)$, which depends on a natural number n. Denote by K the set of numbers n for which $S(n)$ is true. To prove $S(n)$ for all values of n is equivalent to proving that $K = \mathbb{N}$. Peano axiom 5 claims that for this it suffices to prove $S(1)$ (i.e., to prove that $1 \in K$) and to prove implication $S(n) \Rightarrow S(n+1)$ (i.e., together with each $n \in K$ the successor n' of n is also contained in K.)

Exercises

1. Prove that $n' \neq n$ for any $n \in \mathbb{N}$.

2. Prove that the image of map $n \mapsto n'$ is $\mathbb{N} \setminus \{1\}$.

This can be reformulated as follows: for each $n \in \mathbb{N}$ such that $n \neq 1$ there exists $m \in \mathbb{N}$ such that $n = m'$.

3. Prove that the map $\mathbb{N} \to \mathbb{N} \smallsetminus \{1\} : n \mapsto n'$ is invertible.

In other words, the successor of a natural number defines the number uniquely.

The arithmetic operations (addition and multiplication) in the Peano setup are defined axiomatically.

The following two properties define addition:

(1) $n + 1 = n'$

(2) $p + q' = (p + q)'$ for all p and q in $\mathbb{N}$.

These properties can be used to prove associativity and commutativity of addition.

Multiplication is defined by similar properties

(1) $n \cdot 1 = n$

(2) $p \cdot q' = p + (p \cdot q)$.

Peano axioms provide a cumbersome, but logically clean and self-contained approach to natural numbers.